Evaluation of NU-WRF rainfall forecasts for IFloodS

Di Wu[1,2], Christa Peters-Lidard[3], Wei-Kuo Tao[1], and Walter Petersen[4]

1 Mesoscale Atmospheric Processes Laboratory
NASA Goddard Space Flight Center
Greenbelt, Maryland

2 Science Systems and Applications, Inc.
Lanham, Maryland

3 Hydrological Sciences Laboratory
NASA Goddard Space Flight Center
Greenbelt, Maryland

4 Code 610.W
NASA GSFC/Wallops Flight Center
Wallops Island, Virginia

---

[1] *Corresponding author address: Di Wu, Code 612, NASA Goddard Space Flight Center, Greenbelt, MD 20771, Email: di.wu@nasa.gov*

Abstract

The Iowa Flood Studies (IFloodS) campaign was conducted in eastern Iowa as a pre-GPM-launch campaign from 1 May to 15 June 2013. During the campaign period, real time forecasts are conducted utilizing NASA-Unified Weather Research and Forecasting (NU-WRF) model to support the everyday weather briefing. In this study, two sets of the NU-WRF rainfall forecasts are evaluated with Stage IV and Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation (QPE), with the objective to understand the impact of Land Surface initialization on the predicted precipitation. NU-WRF is also compared with North American Mesoscale Forecast System (NAM) 12 km forecast. In general, NU-WRF did a good job at capturing individual precipitation events. NU-WRF is also able to replicate a better rainfall spatial distribution compare with NAM. Further sensitivity tests show that the high-resolution makes a positive impact on rainfall forecast. The two sets of NU-WRF simulations produce very close rainfall characteristics. The Land surface initialization do not show significant impact on short-term rainfall forecast, and it is largely due to the soil conditions during the field campaign period.

1. Introduction

One of the goals of the Global Precipitation Measurement (GPM) mission ground validation program is to conduct integrated hydrological validation, in which the terrestrial water budget is utilized to evaluate the accuracy of blended satellite and/or model-based precipitation products. The Iowa Flood Studies (IFloodS) campaign was conducted in eastern Iowa as a pre-GPM-launch campaign from 1 May to 15 June, 2013, with the goal of examining how well GPM and other blended products could be used for flood forecasting.

The areas of focus for the IFloodS campaign were the Cedar and Iowa River Basins, which were covered by a ground-based NASA NPOL radar deployed along with rain gauges and disdrometers in addition to the existing NEXRAD radar network (Cunha et al., 2015). In addition to the traditional rainfall-oriented instrumentation, in the South Fork Iowa River, a network of rain gauge and soil moisture platforms was deployed in coordination with the Agricultural Research Service and NASA's Soil Moisture Active Passive (SMAP) mission (Coopersmith et al, 2015). These soil moisture platforms, in addition to extensive existing streamflow monitoring by the USGS and Iowa Flood Center, help support the integrated hydrologic validation goals of the campaign. Further, additional, high-resolution integrated hydrologic validation is supported in the Turkey River Basin, within which 20 rain gauges with soil moisture probes and two Iowa Flood Center XPOL weather radars were deployed.

To support deployment of ground-based instrumentation, our team at GSFC conducted real-time forecasting with a meteorological model, which was delivered

daily to support 0900LT forecast briefings delivered to the campaign personnel. This effort required not only dedicated computational resources, but also a robust modeling system capable of simulating severe convective episodes typical of eastern Iowa during the active spring period. In this work we will provide a comprehensive evaluation of the modeling system, including the precipitation forecasts from two different configurations designed to evaluate the impact of land surface initialization on the two forecasts. We first describe the experimental design, including the modeling system, configuration and evaluation datasets. Next, we present an evaluation of the precipitation forecasts based on an archive for the entire experimental period relative to ground data in addition to an operational forecast model. Finally, we discuss the implications of this work for future forecasting applications.

1. Experiment design

a. NU-WRF model

The NASA-Unified WRF (NU-WRF; http://nuwrf.gsfc.nasa.gov) modeling system has been developed at Goddard Space Flight Center (GSFC) as an observation-driven integrated modeling system that represents aerosol, cloud, precipitation and land processes at satellite-resolved scales (Peters-Lidard et al. 2015). NU-WRF is a superset of the National Center for Atmospheric Research (NCAR) Advanced Research WRF (ARW) dynamical core model, achieved by fully integrating the GSFC Land Information System (LIS; Kumar et al. 2006; Peters-Lidard et al. 2015), the WRF/Chem enabled version of the GOddard Chemistry Aerosols Radiation Transport (GOCART; Chin et al. 2000) model, the Goddard Satellite Data Simulation Unit (G-SDSU; Matsui et al. 2009),

98 and custom boundary/initial condition preprocessors. Several NASA physical packages
99 (microphysics and radiation) have also been implemented into NU-WRF. These physical
100 processes include CRM-based microphysics (Tao et al. 2003; Lang et al. 2007, 2011,
101 2014) and radiation (Chou and Suarez 1999) schemes. All the above features are
102 combined into a single software release, with source code available by agreement with
103 NASA/GSFC.

104 In this study, NU-WRF version 3.4.1 (based on NCAR WRF-ARW version 3.4.1)
105 is employed to conduct high-resolution simulations. There are 60 vertical levels and 3
106 spatial domains with 9, 3 and 1 km grid spacing (Fig. 1), and time steps of 27, 9 and 3
107 seconds respectively. The Grell-Devenyi cumulus parameterization scheme (Grell and
108 Devenyi 2002) is adopted for the outer domain; no convective parameterization was used
109 for two inner domains. The PBL parameterization employed the Mellor-Yamada-Janjic
110 (Mellor and Yamada 1982) Level-2 turbulence closure model through the full range of
111 atmospheric turbulent regimes. The Goddard broadband two-stream (upward and
112 downward fluxes) approach was used for the short- and long-wave radiative flux
113 calculations (Chou and Suarez 1999) and its explicit interactions with clouds
114 (microphysics). In addition, the numerical simulations use Goddard 3ICE scheme (Lang
115 et al. 2011), which prognoses three types of ice hydrometeor species (i.e. cloud ice, snow,
116 and graupel).

117

118 **b. NASA Land Information System**

119 The Land Information System (LIS) is a core component of NU-WRF. It is a
120 flexible land surface modeling and data assimilation framework developed with the goal

121 of integrating satellite- and ground-based observational data products and advanced land
122 surface modeling techniques to produce optimal fields of land surface states and physics
123 (Kumar et al. 2006; Peters-Lidard et al. 2007). The infrastructure can not only be
124 directly coupled with the atmosphere, it can also integrate high resolution observations
125 with the model forecasts to generate improved estimates of land surface conditions such
126 as soil moisture, evaporation, snow pack, and runoff, at 1km and finer spatial resolutions
127 and at one-hour and finer temporal resolutions.

128 The role of LIS in the simulation was two-fold: First, to provide physically
129 consistent land surface initialization for NU-WRF; Second, to interact with the surface
130 layer and atmospheric components of NU-WRF and produce coupled water, energy and
131 momentum fluxes. The LSM employed in LIS for this study is Noah LSM version 3.2.1
132 (Ek et al. 2003). WRF-ARW version 3.4.1 uses Noah LSM version 3.4.1. For
133 consistency, the Noah land surface model is run offline within LIS using the same
134 domain configuration as NU-WRF. The Noah LSM in the offline LIS also uses the same
135 soil and vegetation database as NU-WRF.

136 The offline LIS run is cold started from 2 May 2008 to 1 May 2013. The long
137 spin up period is used for land surface states to achieve equilibrium for initialization of
138 WRF-LIS. The LIS offline spin up uses Stage IV rainfall data to provide hourly rainfall,
139 and utilizes GDAS to provide atmospheric forcing input. The Stage IV rainfall product
140 is only used to provide forcing for the LSM during the analysis cycle and is not
141 assimilated into the atmospheric component of the coupled simulation.

142

143 **c. NU-WRF real-time forecasts**

144 During the IFloodS campaign period, two sets of 48-hour NU-WRF forecasts
145 were produced twice a day from May 1st to June 15th 2013. These forecasts required 7
146 hours to produce with 2048 CPUs on NASA NCCS supercomputer.
147 The configuration for both the control (WRF) and coupled (COUP) simulation
148 are as described above (in section 3a). The most substantial difference between the two
149 sets of forecasts are the initialization of the soil states in the land surface model. As
150 previously mentioned, the control simulation uses the same version of Noah Land
151 Surface model as in WRF V3.4.1, while the coupled simulation (COUP) uses a slightly
152 older version (3.2.1) of the Noah Land-Surface model, since that is the most recent
153 version of Noah that was implemented in LIS at the time of the campaign. The changes
154 from 3.2 to 3.4 focused on snow, and based on other offline analyses (not shown) did not
155 result in significant differences in soil moisture, runoff, or land surface fluxes. Hence,
156 the key difference between the WRF and COUP runs is the soil initialization. The soil
157 initialization for the control simulations comes from spatially interpolating the soil
158 moisture and soil temperature states in the forcing dataset, which is NAM in our case,
159 while the land-surface initialization for COUP is produced from a three-year (May 1st
160 2010 to May 1st 2013) offline spin up by LIS, updated daily by an analysis forced by a
161 previous WRF forecast where the precipitation is bias-corrected using the Stage IV
162 blended radar-rain gauge product.
163 As shown in Fig. 2, the forecast starts everyday at 00 UTC and 12 UTC for 48
164 hours integration. For coupled simulation, every forecast cycle starts with a short LIS
165 offline analysis, and then proceeds to an online coupling between WRF and LIS. The
166 offline analysis is initialized from the previous day NU-WRF forecast and forward

integrated for 24 hours to current initialization time. Stage IV data was used to provide hourly precipitation forcing for the LIS analysis, while NU-WRF output from previous day supplements the atmospheric forcing. Similar to the three-year offline spin up, the analysis provides the soil initialization for the coupled simulation, while the atmospheric initialization is provided from NAM, as in the control (WRF) simulation.

d. Data and methodology

Two observational datasets are employed for model evaluation. The Stage IV (Lin and Mitchell 2005) rainfall estimates are available at 4 km grid spacing every hour. It is a widely used rainfall product by both hydrological and meteorological communities, due to its national coverage, high spatial and temporal resolutions and overall low biases (e.g. Tang et al. 2014, Wu et al. 2011, and Seo et al. 2013). Its good performance in Mean Squared Error (MSE) and total bias contributes from the effectiveness of bias correction, and the manual Quality Control (QC) procedures (Cunha et al. 2015).

Multi-Radar/Multi-Sensor System (MRMS) Quantitative Precipitation Estimate (QPE) integrates radar QPE, gauge QPE, local gauge bias corrected radar QPE, and gauge orographic precipitation climatology QPE. MRMS QPE has a 2 min time interval in each 0.01° x 0.01° grid box. One improvement to the previous QPE product 'Q2' is that MRMS uses most advanced dual-polarimetric (DP) radar technologies to eliminate non-meteorological echoes, and it also provides a more accurate spatial distribution of precipitation.

Despite the advantages in DP QPE in certain aspects, it does not necessarily provide an overall superior QPE than single-polarimetric (SP) QPE, such as Stage IV. According to Cunha et al. (2015), DP QPE shows a higher MSE than Stage IV estimates.

190 However, Stage IV also shows a decreased correlation with rain gauges with increasing
191 rainfall threshold (greater than 5 mm/h) than DP estimates. In this study, we choose
192 MRMS as a QPE reference for most statistics analysis. Stage IV is also shown to give a
193 uncertainty range between two datasets.

194 The model analysis is performed on the inner most grid of NU-WRF forecasts,
195 which output each hour with 1 km grid spacing. NAM forecasts are available at three
196 hourly intervals with 12 km spatial resolution. Both 00 Z and 12 Z forecasts are
197 evaluated for 48 hours model integration on each day for all modeling datasets during
198 May 1st to June 15 2015. All datasets are remapped onto the NAM grid and are
199 intercompared at three hourly time intervals.
200

201 **3. Evaluation of rainfall forecast**

202 Figure 3 shows the accumulated precipitation from a composite of NU-WRF
203 simulations and Stage IV dataset from May 1st to June 15th 2013. Stage IV shows clearly
204 a relatively higher accumulated rainfall area from Iowa to northern Illinois compare to
205 surrounding regions. NU-WRF simulated rainfall accumulation shows a similar spatial
206 pattern as the Stage IV. The high accumulation is over Iowa and Northern Illinois, with a
207 peak that has a higher magnitude than Stage IV accumulated rainfall.
208

209 **a. Rainfall time-series**

210 Figure 4 shows three hours accumulated rainfall from two NU-WRF simulations
211 (WRF and COUP), NAM, Stage IV, and MRMS from May 1st to June 15th. The values
212 for models are the mean from available forecast cycles. The first 6 hours simulations are

213 considered as model spin up period, thus are removed from the analysis. Despite the
214 overall overestimation comparing with observations, NU-WRF runs capture the
215 individual precipitation events well. There are only slight differences between the two
216 sets of NU-WRF runs. NAM is very close to observation for averaged rainfall over the
217 whole period. However, it does not replicate the individual events well, especially during
218 May 16th to June 1st, where NAM tends to miss the peak or significantly underestimates
219 the rainfall. During the rest of the periods, NAM has the tendency for overestimation.
220 So despite NAM has a better averaged rainfall over the six-week period than NU-WRF,
221 the NAM does not show a good forecast skill for individual events.
222 Also shown in Fig. 4, the precipitation events come in groups. There are seven
223 wet periods can be identified (Table 1), grouped by at least one dry day (or very light
224 rain) in between two periods. Each precipitation period can be caused by single
225 convective/precipitating system or may be a succession of convective systems. Most of
226 these events have strong upper level support, which is typical for spring and early
227 summer events. Three of these periods have either short wave troughs (VII) or a
228 combination of short wave and long wave troughs (III and IV), which brings weak but
229 complicated forcings to the region. Especially for period IV, where one short wave
230 comes after another, and it associates with a series of propagating systems at surface.
231 NAM struggles to produce an accurate precipitation forecast during period IV
232 (Fig. 4). On the other hand, NU-WRF is able to reproduce the individual peaks during
233 period IV. Despite using NAM as initial and boundary condition, the finer resolution
234 NU-WRF simulations certainly show improved features in reproducing these
235 precipitation events. These series of weakly forced events demonstrate the characteristics

236 of warm season MCSs. And according to many studies (e.g. Doswell et al. 1996; Fritsch
237 and Carbone 2004), it is very difficult to improve forecasts for deterministic warm season
238 rainfall events. Even with the advancements in numerical models in recent years in both
239 physics and resolutions, the forecast for warm season convections still remain a
240 challenge. And since it brings most heavy rainfall and events with high societal impacts,
241 it is also quite important to improve warm season deterministic rainfall forecast.
242
243 **b. Rainfall statistics**
244 The domain averaged three hours accumulated rainfall statistical scores are
245 calculated from May 1st to June 15th with respect to the forecast lead-time (Fig. 5). The
246 bias scores (Fig. 5a) are negative for WRF, COUP, and NAM in the first six hours, which
247 is due to the cold start of these models, where all the precipitation values are initialized
248 from zero. The two NU-WRF simulations have appreciably higher biases scores than
249 NAM comparing with MRMS, which is also evident from Fig. 4 that NAM
250 underestimates rainfall during period III and IV while overestimates for other periods.
251 Despite the lower bias, NAM has higher RMSE (0.73) than two NU-WRF simulations
252 (0.67 and 0.69). Also NAM has lower correlation (0.60) comparing with two NUM-
253 WRF simulations (0.76 and 0.73). Both NU-WRF domain averaged rainfall has
254 relatively high correlations with MRMS, which echoes that NU-WRF has captured
255 individual precipitation events in Fig. 4. The correlation trend between models and
256 MRMS are only decreasing slightly with increase of forecast lead-time. The correlation
257 between Stage IV and MRMS are very high (0.98). The differences between two NU-
258 WRF simulations are small enough comparing with the differences between Stage IV and

MRMS. However, the differences between two NU-WRF correlations increase with forecast lead-time.

When considering spatial variability between different datasets and MRMS, the correlation scores are much lower than the area averaged quantities (Fig. 6). The correlation between Stage IV and MRMS decreases from 0.98 for area-averaged rainfall to 0.82 for considering both time and spatial correlations. The score drops significantly for correlations between models and MRMRS, from previously 0.7 to under 0.2. From Fig. 6, the decreasing trend of correlation is obvious for models, which demonstrate the forecast skill has decreased with increase of forecast hours. Despite the low spatial correlation, NU-WRF seems to produce consistently slightly higher correlation than NAM forecasts.

Table 2 shows the spatial correlation scores for all seven different forecast periods. NAM shows consistently lower scores than both NU-WRF forecasts. And the correlation scores vary with different periods. NAM has the lowest correlation during period IV and VII, while NU-WRF has a relatively lower correlation during V and VII. Period IV is one difficult period for NAM, on the other hand, NU-WRF did a fine job at capturing individual events (not shown) during this period. It is one of the periods that NAM and NU-WRF has the largest differences in correlation score (0.07 vs 0.18).

One caveat of spatial correlations analysis is that the displacement in spatial correlations is heavily penalized, same as the appearance of spurious precipitating regions. Thus for individual time slice, the correlation may not necessarily reflect all aspects of the forecast performance. Such as in Fig. 7, even the NU-WRF has captured the characteristics of heavy precipitation, it still has a lower correlation score. However,

282 with a relatively bigger sample size, the correlation is capable of describing a general
283 trend. Such as, NAM consistently have a lower spatial correlation score than NU-WRF
284 (Fig. 6), which is consistent with that NAM has a lower domain averaged correlation as
285 well (Fig. 5c). On the other hand, identification for spatial displacement is also
286 important. It is the dominant source of quantitative precipitation forecast (QPF) error
287 (Ebert and McBride 2000). Poor QPF skill has hindered hydrologic applications,
288 particularly streamflow forecasting operations (Cuo et al. 2011).

289 The two NU-WRF runs have very close correlations with each other, and the
290 differences grow with time, which is also observed from rainfall spatial distributions for
291 various cases (not shown). It is a question whether the spread is caused by physical
292 differences between the two models or by random error growth. From Fig. 6a, the
293 differences are fluctuating after 12 hours into the forecast for all cases, and similarly for
294 different periods. However, a systematic evaluation of this error behavior is beyond the
295 scope of this study.

296 Figure 8 shows rainfall statistics for each forecast cycle using MRMS as a
297 reference. So the value at May 15th shows correlation of MRMS and 6 to 48 hours model
298 forecast initialized at 00 UTC on May 15th. The overall scores are similar to the scores
299 from Fig. 5. The differences are due to the sample selection and whether or not including
300 the first six hours into the considerations. In addition to aforementioned differences
301 between models, it also clearly shows how models perform during different periods. Fig.
302 8a shows models underestimate some rainfall events during period III, IV, and VII, while
303 overestimate the rest of the periods. NAM shows large negative bias during period IV
304 and positive bias during other periods, so the overall low bias of NAM is merely an

305 averaged effect from positive and negative biased cases. For the RMSE score, all models
306 show similar score (~0.55) compare with MRMS. The spread of the RMSE scores
307 among models (0.02) are even smaller than that between Stage IV and MRMS (0.14).
308 High model RMSE are also associated with heave precipitation periods. Period III and
309 IV produce many heavy precipitation events, also associated with high model RMSEs.
310 Figure 8c shows the correlations between different datasets and MRMS. NAM
311 performed poorly during period III and IV, which is also reflected in Fig. 4. The low
312 correlation period is around May 11th to 13th, where not much rainfall was brought to the
313 area (Fig. 4). Small phase shift in rainfall measurements can induce low correlation
314 scores between Stage IV and MRMS. Models also have poor correlations with MRMS
315 during this period. The differences of correlation scores of the two NU-WRF
316 simulations (0.01) are smaller than that of Stage IV and MRMS (0.05).
317
318 **c. Rainfall time-series PDF**
319 The precipitation PDF is also evaluated for model and observational datasets. As
320 shown on Fig. 9, NU-WRF compares very well with Stage IV data for rainfall
321 accumulations less than 8 mm every three hours, but there are some overestimations for
322 heavier rainfall frequencies. On the other hand, NAM has produced outstandingly high
323 frequencies for very light precipitation (0.25-0.5 mm), while significantly underestimated
324 the frequencies for heavy rainfall (> 4 mm). This feature is also evident from rainfall
325 spatial distribution (Fig. 7), NAM produces large light rain area, but fails to produce any
326 heavy rainfall. It is also quite common in many coarse resolution models; large grid
327 spacing limits its ability to resolve sub-grid convections.

328 The two sets of NU-WRF runs have very close PDF distribution. The differences
329 of the two NU-WRF runs are even smaller than those between the two observational
330 dataset (Stage IV and MRMS). MRMS has higher light rainfall (0.25-1 mm) coverage
331 and higher coverage for intense rainfall (>16 mm). The small differences between two
332 NU-WRF runs indicate that different land surface initialization and differences within the
333 LSM do not have a big impact on precipitation intensity. On the other hand, the PDF
334 distribution is shown to be more sensitive to different cloud microphysics schemes (Tao
335 et al., 2015).
336
337 **4 Discussion**
338 **a. Sensitivity Tests**
339 NU-WRF and NAM share the same initial and boundary conditions, but their
340 predicted rainfall characteristics are very different (shown in section 3), especially during
341 period IV. Additional sensitivity tests are performed for 9 km (WRF_9km) and 3km
342 (WRF_3km) resolution, using 1 km NU-WRF without LIS coupling (WRF) as control
343 run. WRF_9km employs only the outer most domain in Fig. 1, while the WRF_3km uses
344 two outer domains in Fig.1. Besides with different resolutions and domain setups,
345 WRF_9km uses Grell-Devenyi ensemble cumulus scheme (GD), where the WRF and
346 WRF_3 km runs only applied GD for the outer most grid (with 9 km resolution). NAM
347 has significantly underestimated rainfall and the peak is out of phase with the
348 observations for forecast initiated at 00 UCT on May 29th 2013 (Fig. 10). WRF and
349 COUP, on the other hand, captured the peak, despite 3 hours delayed than the
350 observation. By reducing the resolution from original 1 km (WRF) to 9 km (WRF_9km),

351 there is a reduction of peak value and shift of phase. However, the change of resolution
352 from 1 km to 3 km (WRF_3km) does not result in significant changes in the forecast. In
353 addition, 9 km run with Betts-Miller-Janjic (BMJ) cumulus scheme (WRF_BMJ) is used
354 to compare with 9 km with the GD cumulus scheme (WRF_9km), same GD scheme was
355 applied to the outer most domain in the control run (WRF). BMJ cumulus scheme is also
356 the one used in NAM forecast. WRF_BMJ has an even lower peak magnitude than
357 WRF_9km run, and its averaged rainfall is the most comparable to NAM in all the
358 simulations. There is improved forecast skill by adopting finer resolution and using GD
359 cumulus scheme, even when initial and boundary conditions stay the same. This is also
360 consistent with many previous studies (Wang and Seaman 1997; Gallus 1999), the choice
361 of convective schemes have a strong influence on simulated rainfall pattern.
362
363 **b. Soil moisture and rainfall**
364 Of particular interest in this study is that whether there are improvements in
365 rainfall forecast by applying high resolution and more accurate land-surface initial
366 condition comparing to interpolated fields from regional model forecast. As shown in
367 previous sections, the differences in the rainfall forecasts between the two NU-WRF runs
368 are rather small. One possible reason for such small differences is that the region of
369 interest is under the influence of many heavy precipitation events during the campaign
370 period, thus with high water availability, the moisture transport from surface are similarly
371 high in both models. One indication for the above argument is high evaporative fraction
372 (EF), which is the ratio of latent heat to available energy at the land surface. EF is a
373 diagnostic for the surface energy balance (energy-limited state or a moisture-limited

374 state), supposedly isolates soil moisture and vegetation from radiation and turbulent
375 factors. Despite the strong diurnal periodicity for two components in the surface energy
376 balance, EF is generally considered to be a constant during daytime hours (Nichols and
377 Cuenca 1993; Crago 1996; Crago and Brutsaert 1996). Figure 11 shows the daily EF
378 from NU-WRF run, which is averaged from local 7 am to 6 pm. EF stays over 0.6 for
379 most days, which means the energy fluxes to the surface energy budget are mainly
380 contributed by latent heating. With high EF, the impact for precipitation processes with
381 different soil initialization is minimized. Figure 12 shows the evaporative fraction at
382 local noon for NU-WRF forecast initialized at 12 UTC on May 30th 2013, and it also
383 shows the 0-10 cm soil moisture at model initialization. WRF has higher soil moisture
384 than COUP (Figure 12 and 13), but COUP has an only slightly smaller EF than WRF.
385 The slightly dryer top soil moisture from COUP is also observed in Goergia and South
386 Carolina during summer season, which is actually closer to the observed soil moisture
387 from the U.S. Department of Agriculture's Soil Climate Analysis Network (SCAN) (Case
388 et al. 2011). The small EF difference shows that their similar partition in surface energy
389 budget. WRF also has a much lower resolution than COUP. The lower resolution is due
390 to the interpolated soil moisture field from NAM that has a 12 km resolution, while
391 COUP uses LIS offline spin up that provides soil moisture at a resolution that is
392 consistent with NU-WRF grid. In addition, COUP uses Stage IV observed rainfall to
393 force the LIS offline spin up, which should result a more accurate and observational
394 consistent soil moisture profile than WRF. Despite with fine resolution and more
395 accurate soil moisture initialization, the high EF indicates COUP and WRF have a similar

land surface moisture transport to atmospheric boundary layer, which will minimize their impact on precipitation processes.

5. Conclusions

Two sets of NU-WRF are used for providing real-time forecasts twice a day for two-day long integration during IFloodS field campaign from May 1st to June 15th. One of the NU-WRF forecast uses NAM interpolated land surface field as LSM forcing; the other one uses LIS spin up to provide land surface conditions, which assimilates the latest Stage IV observed precipitation. These two sets of model datasets are compared with low resolution forcing dataset (NAM) and with each other. The precipitation fields are evaluated with Stage IV and MRMS. Two observations datasets are able to provide a difference range, which indicates observational uncertainties. The main conclusions are as follows:

1) Both NU-WRF simulations are able to reproduce the individual precipitation event during the field campaign period, which NAM is out of sync with observations for heavy precipitation events during period IV. In addition, for those events where rainfall intensity less than 1 mm/h, NAM tends to overestimates the rainfall amount. However, for heavy rainfall events (e.g. May 20th, May 25th to May 30th, and June 13th), NAM has underestimated the rainfall amounts. Despite NAM has a better averaged rainfall over the six-week period compared with NU-WRF, the NAM forecast skill is not necessarily better for individual events. Even though NU-WRF shares the same initial and boundary condition from NAM, and its rainfall field is averaged onto the coarser NAM grid, NU-WRF still outperforms NAM in both time and spatial correlations.

2) NU-WRF is also able to produce a better rainfall PDF than NAM. NAM significantly underestimates the frequencies for heavy rainfall and largely overestimates frequencies for very light rainfall. While NU-WRF is able to produce PDF that is very close to the observed distribution from Stage IV.

3) NU-WRF sensitivity tests show that by switching to a coarser resolution and to a different convective scheme, the rainfall forecast skill has reduced, which turns out to be more comparable to NAM forecast in a case study. Model resolution makes a difference for rainfall forecast, but it also depends on which scale it applies to, whether within the convective permitting scale or to coarser scales.

4) LIS spin up with Stage IV forcing has the advantage of producing higher resolution and more accurate surface properties than without LIS spin up. However, the benefit for the precipitation forecast is marginal. Two sets of NU-WRF simulations do not yield significant differences on rainfall characteristics during IFloodS field campaign period. The differences between two NU-WRFs are much smaller than the differences between NAM and NU-WRF or between the two observational datasets (Stage IV and MRMS). Evaporative fraction indicates the relatively similar land surface moisture transport between the two NU-WRF simulations, which inhibit the land-surface improvement to have a positive impact on precipitation forecast.

Acknowledgements

This research was supported by the NASA Precipitation Measurement Missions (PMM) and Modeling and Analysis Program (MAP) solicitations through awards to PI's Peters-Lidard, Tao and Petersen. This support is gratefully acknowledged. We also

442 acknowledge the excellent computational and storage support provided by NASA's
443 Center for Climate Simulation (NCCS).
444

445

446 **References**

447 Carbone, R. E., J. D. Tuttle , D. A. Ahijevych , and S. B. Trier, 2002: Inferences of
448 Predictability Associated with Warm Season Precipitation Episodes. *J. Atmos.*
449 *Sci*., **59**, 13

450 Case, J. L., S. V. Kumar, J. Srikishen, and G. J. Jedlovec, 2011: Improving Numerical
451 Weather Predictions of Summertime Precipitation over the Southeastern United
452 States through a High-Resolution Initialization of the Surface State. Wea.
453 Forecasting, **26**, 785–807. doi:http://dx.doi.org/10.1175/2011WAF2222455.1

454 Chin, M., R. B. Rood, S.-J. Lin, J. F. Muller, and A. M. Thomspon, 2000: Atmospheric
455 sulfur cycle in the global model GOCART: Model description and global
456 properties, *J. Geophys. Res*., **105**, 24,671-24,687.

457 Chou, M.-D., and M. J. Suarez, 1999: A shortwave radiation parameterization for
458 atmospheric studies, *NASA/TM-104606*, **15,** pp 40.

459 Coopersmith, Evan J., Michael H. Cosh, Walt A. Petersen, John Prueger, and James J.
460 Niemeier, 2015: Soil Moisture Model Calibration and Validation: An ARS
461 Watershed on the South Fork Iowa River. *J. Hydrometeor*., **16**, 1087–1101. doi:
462 http://dx.doi.org/10.1175/JHM-D-14-0145.1.

463 Crago, R., 1996: Conservation and variability of the evaporative fraction during the
464 daytime. *J. Hydrol.,* **180**, 173–194.

465 Crago, R., and W. Brutsaert, 1996: Daytime evaporation and the self-preservation of the
466 evaporative fraction and the Bowen ratio. *J. Hydrol.*, **178**, 241–255.

Cunha, Luciana K., James A. Smith, Witold F. Krajewski, Mary Lynn Baeck, Bong-Chul Seo, 2015: NEXRAD NWS Polarimetric Precipitation Product Evaluation for IFloodS. *J. Hydrometeor*, Early online release, doi: http://dx.doi.org/10.1175/JHM-D-14-0148.1

Cuo, L., T. C. Pagano, and Q. J. Wang, 2011: A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *J. hydrometeor*., **12**, 713-729

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res*., **108**(D22), 8851, doi:10.1029/2002JD003296

Gallus, W. A., Jr., 1999: Eta simulations of three extreme rainfall events: Impact of resolution and choice of convective scheme. *Wea. Forecasting*, **14**, 405–426.

Gallus, W. A., and M. Segal, 2000: Sensitivity of forecast rainfall in a Texas convective system to soil moisture and convective scheme. *Wea. Forecasting*, **15**, 509–526.

Grell, G. A., and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophy. Res. Lett.*, **29,** Article 1693.

Kumar, S. V., Y. Tian, C. Peters-Lidard, and Coauthors, 2006: Land information system: An interoperable framework for high resolution land surface modeling. *Environ. Modelling Software,* **21,** 1402-1415.

488 Lang, S., W.-K. Tao, R. Cifelli, W. Olson, J. Halverson, S. Rutledge, and J. Simpson,
489 2007: Improving simulations of convective system from TRMM LBA: Easterly
490 and Westerly regimes, *J. Atmos. Sci*., **64**, 1141-1164.
491 Lang, S. E., W.-K. Tao, X. Zeng, and Y. Li, 2011: Reducing the biases in simulated radar
492 reflectivities from a bulk microphysics scheme: Tropical convective systems, *J.*
493 *Atmos. Sci*., **68**, 2306–2320.
494 Lang, S., W.-K. Tao, J.-D. Chern, D. Wu, and X. Li, 2014: Benefits of a 4th ice class in
495 the simulated radar reflectivities of convective systems using a bulk microphysics
496 scheme, *J. Atmos. Sci*.,**71**, 3583-3612.
497 Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses:
498 Development and application. Preprints, *19th Conf. on Hydrology, Am. Meteorol.*
499 *Soc.*, San Diego, CA, P1-2.
500 Matsui, T., X. Zeng, W.-K. Tao, H. Masunaga, W. S. Olson, and S. Lang, 2009:
501 Evaluation of long-term cloud-resolving model simulations using satellite
502 radiance observations and multi-frequency satellite simulators. *J. Atmos. Oce.*
503 *Tech.,* **26**, 1261-1274.
504 Matsui, T., J. Santanello, J. J. Shi, W.-K. Tao, D. Wu, C. Peters-Lidard, E. Kemp, M.
505 Chin, D. Starr, M. Sekiguchi, and F. Aires, 2014: Introducing Multi-Sensor
506 Satellite Radiance-based Evaluation for Regional Earth System Modeling, *J.*
507 *Geophy. Res.*, **119**, 8450-8475, doi: http://dx.doi.org/10.1002/2013JD021424
508 Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for
509 geophysical fluid problems, *Rev. Geophys. Space Phys*., **20**, 851-875.

Nichols, W. E., and R. H. Cuenca, 1993: Evaluation of the evaporative fraction for parameterization of the surface energy balance. *Water Resour. Res.,* **29**, 3681–3690.

Peters-Lidard, C.D., E. M. Kemp, T. Matsui, J. A. Santanello, Jr., S. V., Kumar, J. P. Jacob, T. Clune, W.-K. Tao, M. Chin, A. Hou, J. L. Case,, D. Kim, K.-M. Kim, W. Lau, Y. Liu, J.-J. Shi, D. Starr, Q. Tan,, Z. Tao, B. F. Zaitchik, B. Zavodsky, S. Q. Zhang, M. Zupanski (2015), Integrated modeling of aerosol, cloud, precipitation and land processes at satellite-resolved scales, *Environmental Modelling & Software*, **67**, 149–159. doi:http://dx.doi.org/10.1016/j.envsoft.2015.01.007

Seo, B.-C., L. K. Cunha, and W. F. Krajewski, 2013: Uncertainty in radar- rainfall composite and its impact on hydrologic prediction for the eastern Iowa flood of 2008. *Water Resour. Res*, **49**, 2747–2764.

Tang, L., Y. Tian, and X. Lin, 2014: Validation of precipitation retrievals over land from satellite-based passive microwave sensors. *J. Geophys. Res. Atmos*. **119** (**8**): 4546-4567 [10.1002/2013JD020933]

Tao, W.-K., J. Simpson, D. Baker, S. Braun, M.-D. Chou, B. Ferrier, D. Johnson, A. Khain, S. Lang, B. Lynn, C.-L. Shie, D. Starr, C.-H. Sui, Y. Wang and P. Wetzel, 2003: Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GCE) model, A Special Issue on Non-hydrostatic Mesoscale Modeling, *Meteor. Atmos. Phys*., **82**, 97-137.

Tao, W. K., D. Wu, S. Lang, J. Chern, C. Peters-Lidard, A. Fridlind, and T. Matsui, 2015: High-resolution NU-WRF simulations of a deep convective-precipitation system

during MC3E: Part I: Comparisons between Goddard microphysics schemes and observations. *J. Geophys. Res. Atmos.,* in revision.

Wang, W., and N. L. Seaman, 1997: A comparison study of convective schemes in a mesoscale model. *Mon. Wea*. *Rev*., **125**, 252–278.

Wu, W., D. Kitzmiller, and S. Wu, 2011: Evaluation of radar precipitation estimates from the National Mosaic and Quantitative Precipitation Estimation System and the WSR-88D Precipitation Processing System over the Conterminous United States. *J. Hydrometor*., **13**, 1080–1093.

Period	Dates	Synoptic
I	5/2-5/5	Trough and surface front
II	5/8-5/11	Low
III	5/16-5/24	Short wave trough followed by long wave trough
IV	5/25-6/2	Group of short wave trough
V	6/4-6/7	Low
VI	6/8-6/11	Trough
VII	6/12-6/14	Two short wave troughs

Table 1. Seven precipitation periods and their synoptic setups.

Period	I	II	III	IV	V	VI	VII
WRF	0.22	0.20	0.15	0.18	0.11	0.16	0.09
COUP	0.22	0.21	0.15	0.18	0.11	0.16	0.12
NAM	0.20	0.17	0.10	0.07	0.09	0.13	0.07

546 Table 2. Rainfall spatial correlations between model forecasts with MRMS during the
547 seven precipitation periods.
548
549 **List of Figures**
550 Figure 1: NU-WRF grid configuration. The outer domain (labeled 1 at the center) has a
551 horizontal resolution of 9 km. The middle domain (labeled 2) has a horizontal resolution
552 of 3 km, and the inner domain (labeled 3) has a horizontal resolution of 1 km and covers
553 Iowa.
554 Figure 2: Flow chart for real-time forecast using NU-WRF coupling with LIS. Orange
555 boxes show the initial time of forecast cycles. Green boxes show LIS offline spin up to
556 provide soil initial condition. The LIS spin up uses Stage IV to provide precipitation
557 forcing and NU-WRF previous day forecast to provide additional meteorological forcing.
558 Blue boxes show NU-WRF coupling with LIS uses LIS offline spin up to provide surface
559 initial conditions and use NAM to provide meteorological initial conditions and boundary
560 conditions.
561 Figure 3: Accumulated precipitation from NU-WRF real-time forecast and Stage IV
562 dataset from May 1st to June 15th 2013.
563 Figure 4: Three hours accumulated precipitation from NU-WRF with LIS (CP) and
564 without LIS coupling (WRF), NAM, MRMS, and Stage IV datasets from May 1st to June
565 15th 2013.
566 Figure 5: Time series of bias (a), root mean square error (b), and correlation (c) of
567 domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against
568 MRMS with respect to forecast hours.

569 Figure 6: Similar to Figure 5 (c), but for spatial correlations between NU-WRF and NAM
570 forecasts with MRMS for the whole campaign period from May 1st to June 15th (a), for
571 period III May 16th to 24th(b), and for period IV May 25th to June 2nd (c). Stage IV has a
572 0.82 correlation averaged through out the campaign period.
573 Figure 7: Three hours rainfall accumulation (mm) for MRMS, NAM, NU-WRF with LIS
574 coupling (COUP) and without LIS (WRF) at 06 UTC on May 20th 2013, which is 18
575 hours since model initialization.
576 Figure 8: Time series of bias (a), root mean square error (RMSE) (b), and correlation (c)
577 of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against
578 MRMS with respect to each forecast.
579 Figure 9: PDF of observed and forecasted three hours accumulated precipitation from
580 May 1st to June 15th 2013.
581 Figure 10: Domain averaged three hours rainfall accumulation for NU-WRF sensitivity
582 runs. The result is from forecast cycle initiated at 00 UTC on May 29th, 2013.
583 Figure 11: Daily averaged (7 am to 6 pm) evaporative fraction from May 1st to June 15th
584 2013.
585 Figure 12: Evaporative fraction at local noon (17 UTC on May 19th) and top soil moisture
586 (0-10 cm) at model initialization, which is at 12 UTC on May 19th, 2013.
587 Figure 13: Same as Figure 11, except for daily averaged soil moisture.
588
589

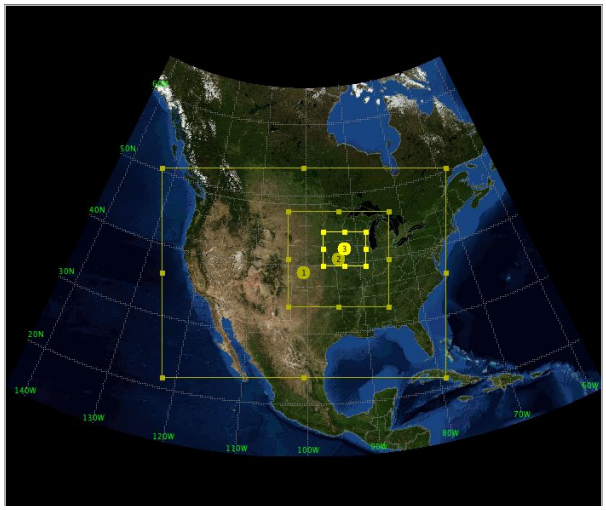


Figure 1: NU-WRF grid configuration. The outer domain (labeled 1 at the center) has a horizontal resolution of 9 km. The middle domain (labeled 2) has a horizontal resolution of 3 km, and the inner domain (labeled 3) has a horizontal resolution of 1 km and covers Iowa.

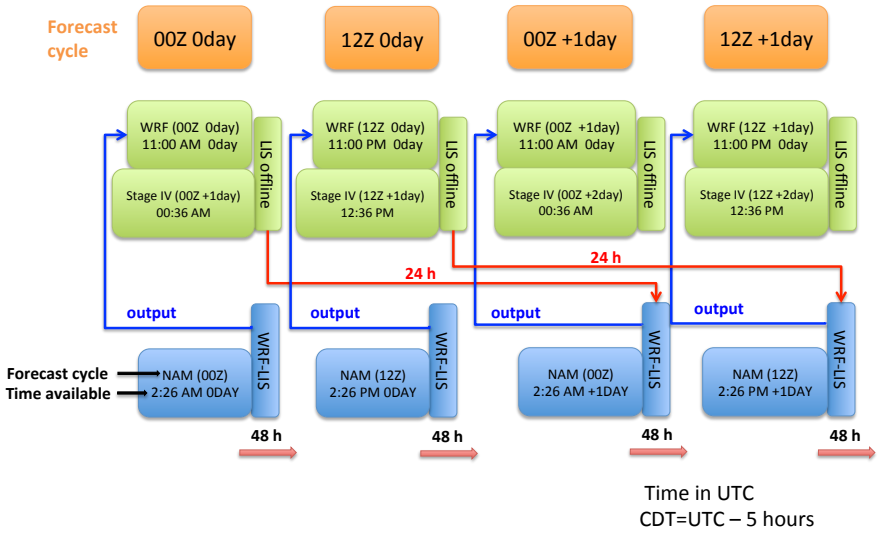


Figure 2: Flow chart for real-time forecast using NU-WRF coupling with LIS. Orange boxes show the initial time of forecast cycles. Green boxes show LIS offline spin up to

provide soil initial condition. The LIS spin up uses Stage IV to provide precipitation forcing and NU-WRF previous day forecast to provide additional meteorological forcing. Blue boxes show NU-WRF coupling with LIS uses LIS offline spin up to provide surface initial conditions and use NAM to provide meteorological initial conditions and boundary conditions.

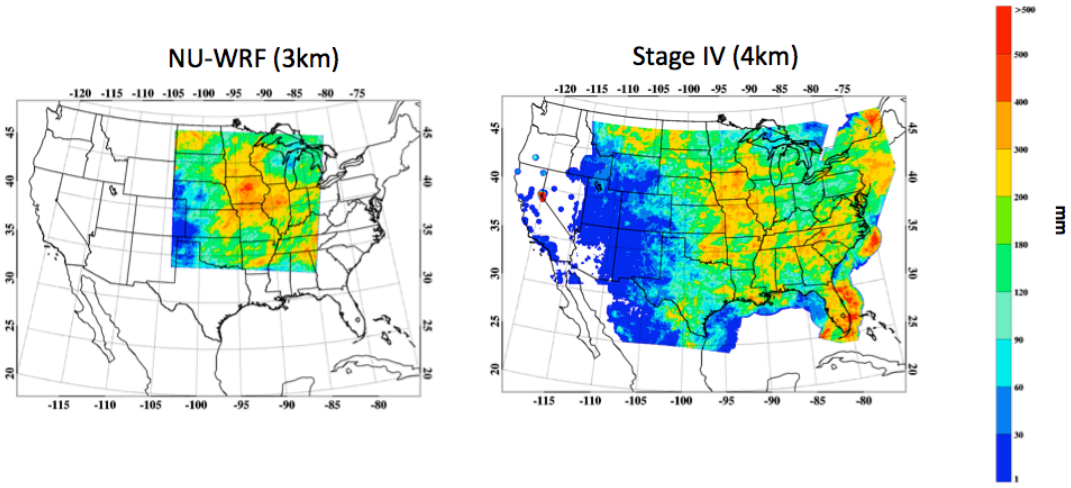
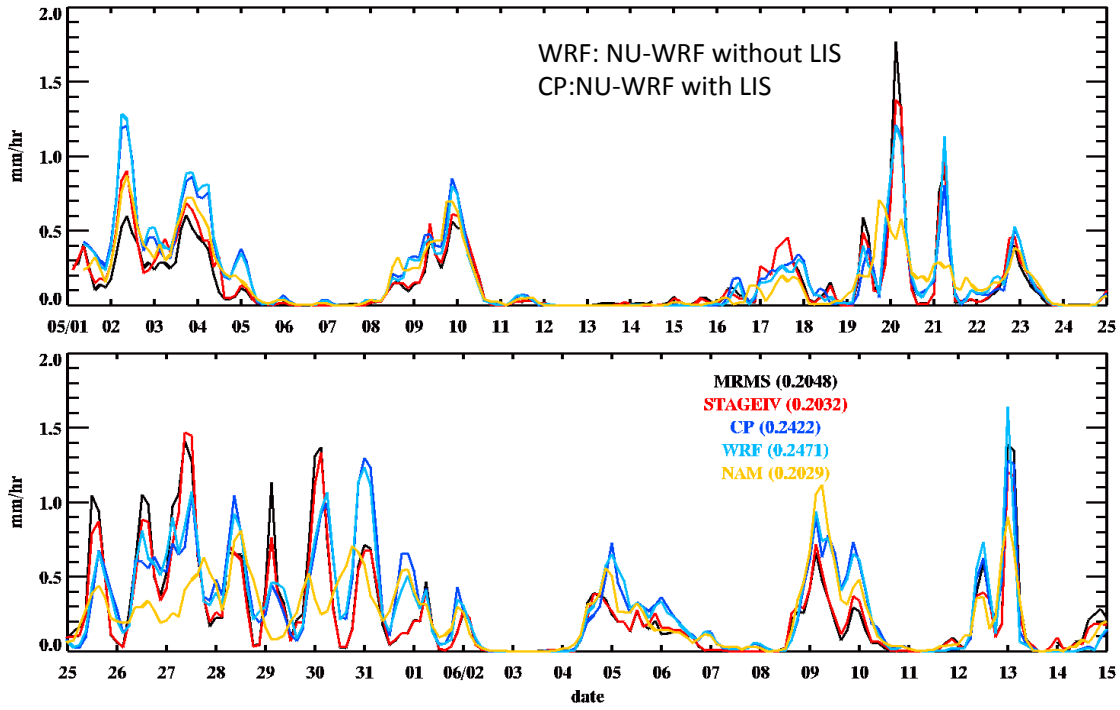


Figure 3: Accumulated precipitation from NU-WRF real-time forecast and Stage IV dataset from May 1st to June 15th 2013.



612

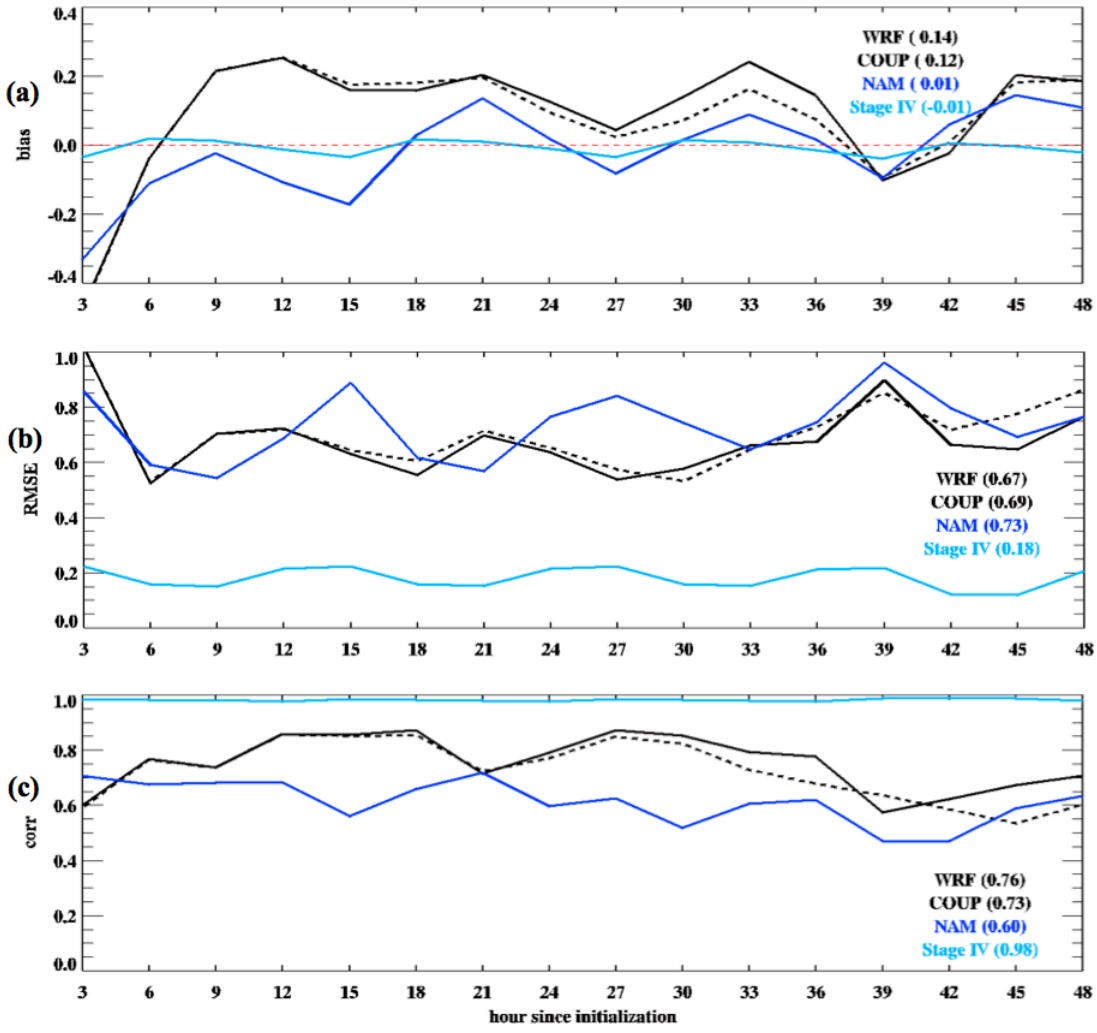
613

614 Figure 4: Three hours accumulated precipitation from NU-WRF with LIS (CP) and

615 without LIS coupling (WRF), NAM, MRMS, and Stage IV datasets from May 1st to June

616 15th 2013.

617



618

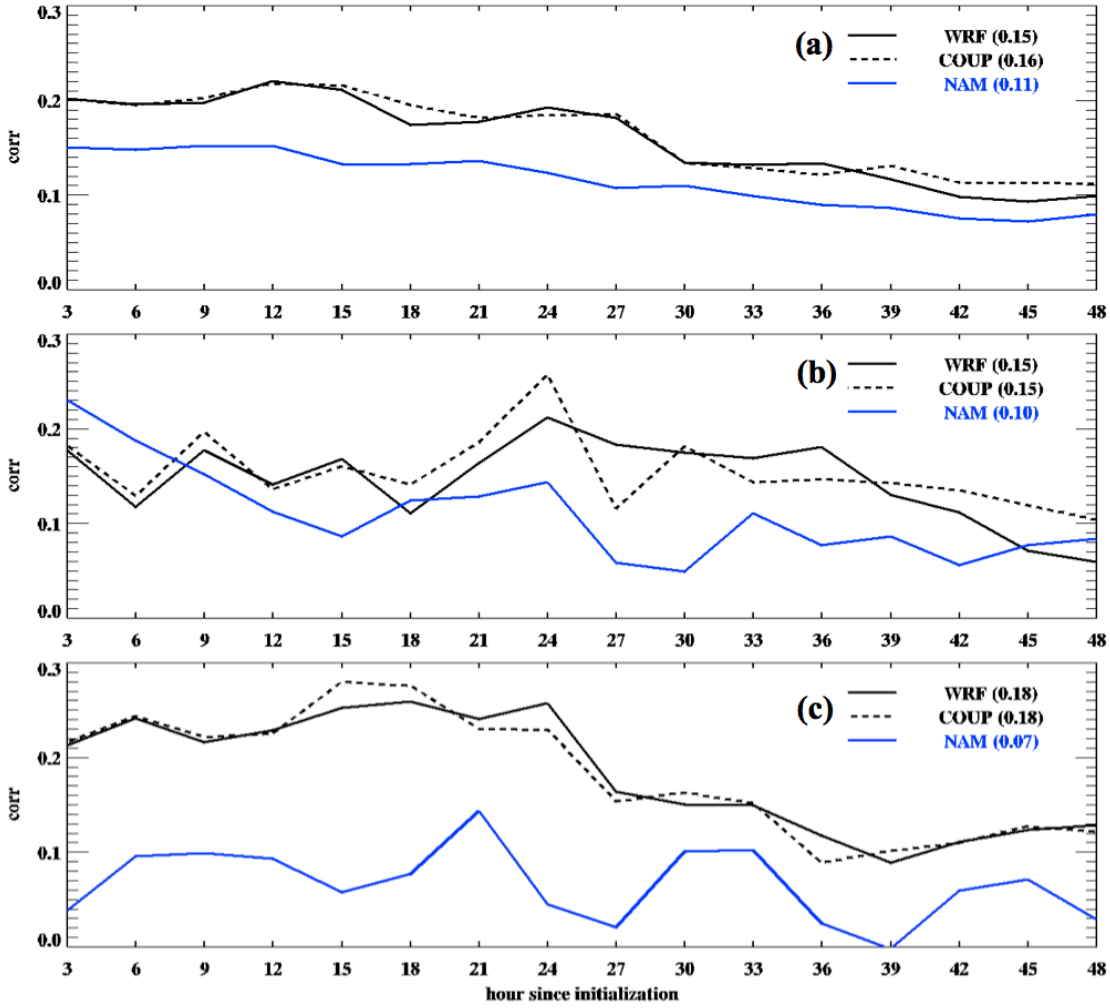
619

620 Figure 5: Time series of bias (a), root mean square error (b), and correlation (c) of

621 domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against

622 MRMS with respect to forecast hours.

623



624

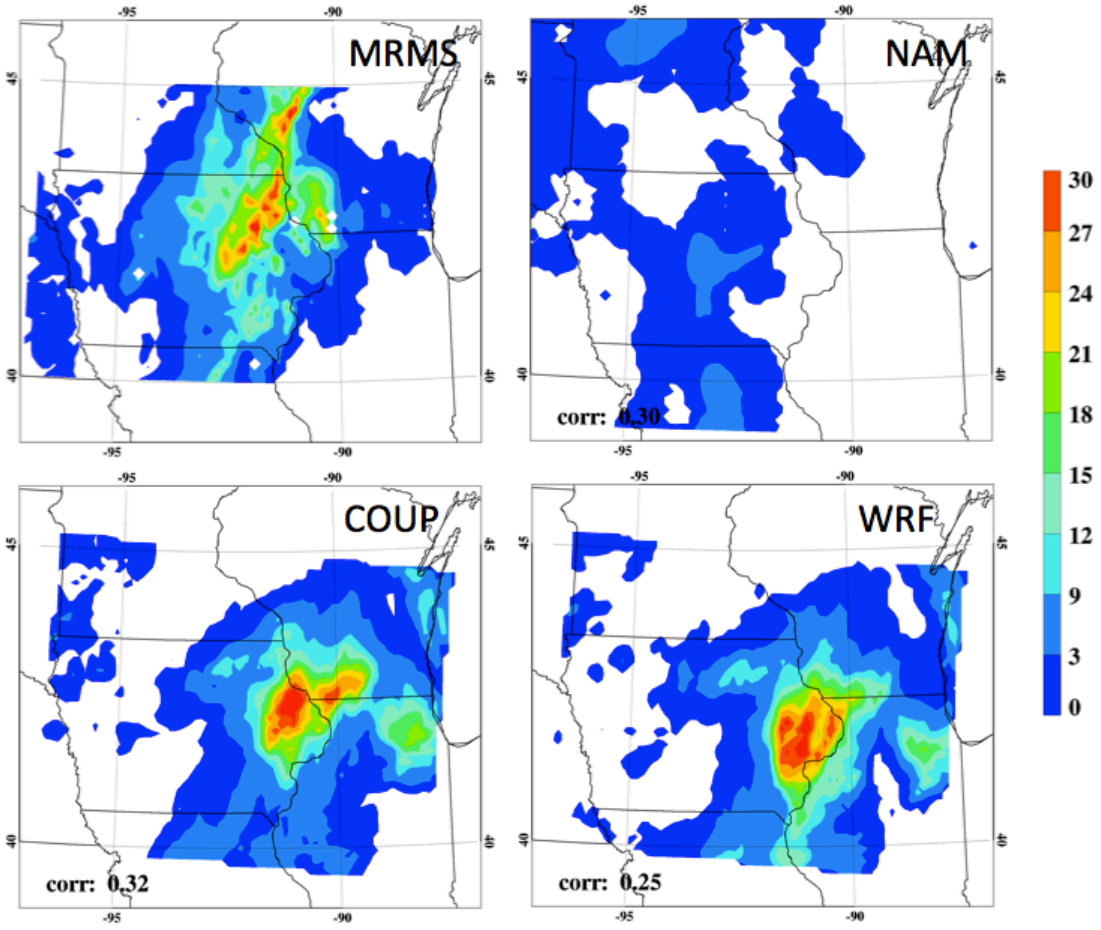
625 Figure 6: Similar to Figure 5 (c), but for spatial correlations between NU-WRF and NAM

626 forecasts with MRMS for the whole campaign period from May 1st to June 15th (a), for

627 period III May 16th to 24th(b), and for period IV May 25th to June 2nd (c). Stage IV has a

628 0.82 correlation averaged through out the campaign period.

629



630

631 Figure 7: Three hours rainfall accumulation (mm) for MRMS, NAM, NU-WRF with LIS

632 coupling (COUP) and without LIS (WRF) at 06 UTC on May 20th 2013, which is 18

633 hours since model initialization.

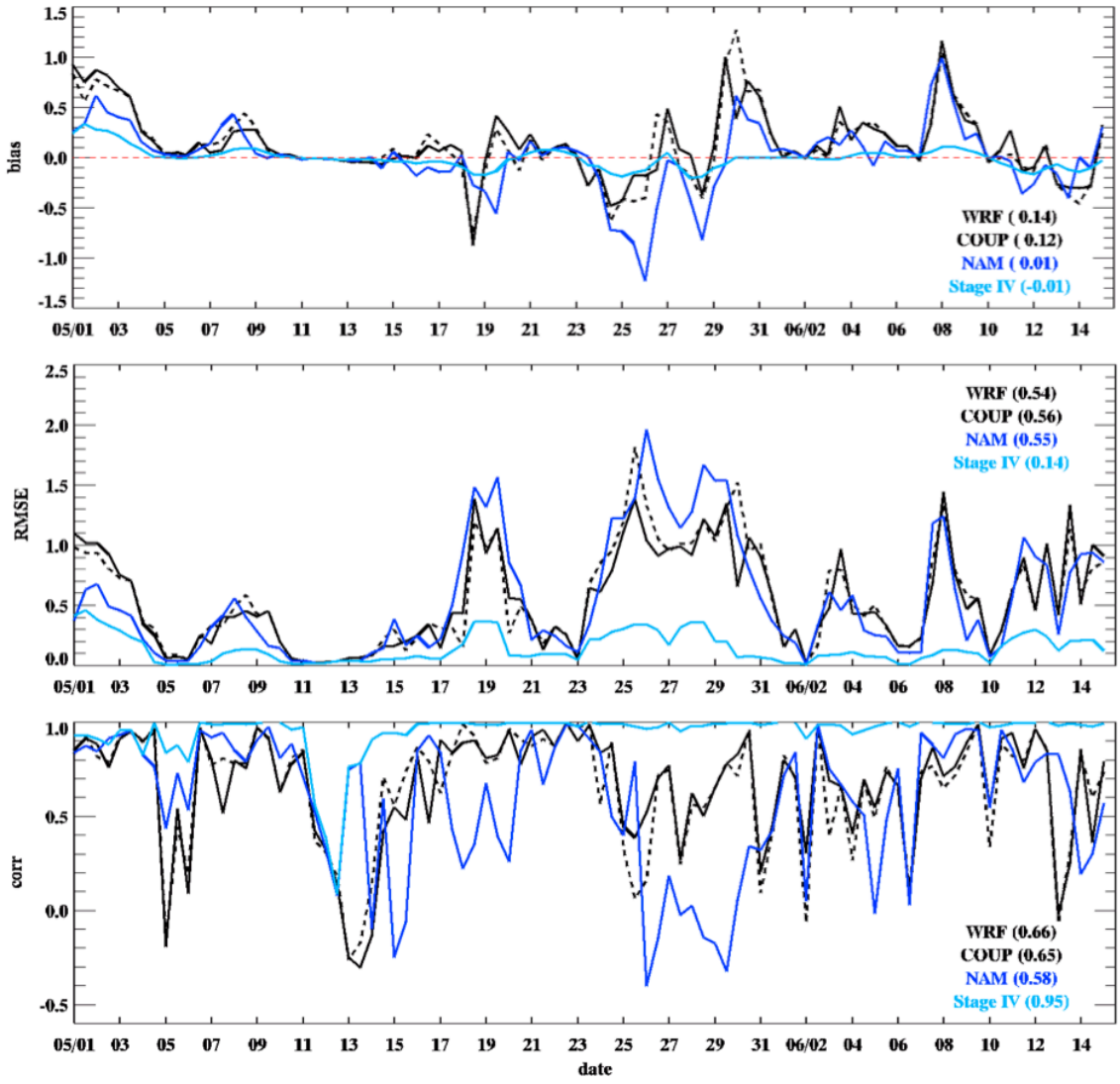
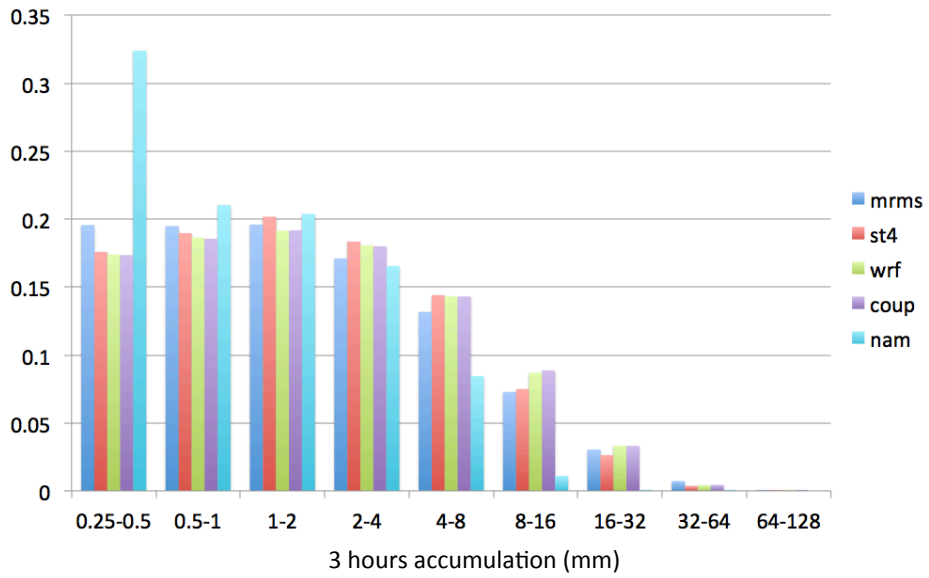


Figure 8: Time series of bias (a), root mean square error (RMSE) (b), and correlation (c) of domain averaged rainfall from WRF, COUP, NAM, and Stage IV compared against MRMS with respect to each forecast.



641

642 Figure 9: PDF of observed and forecasted three hours accumulated precipitation from

643 May 1st to June 15th 2013.

644

645

646

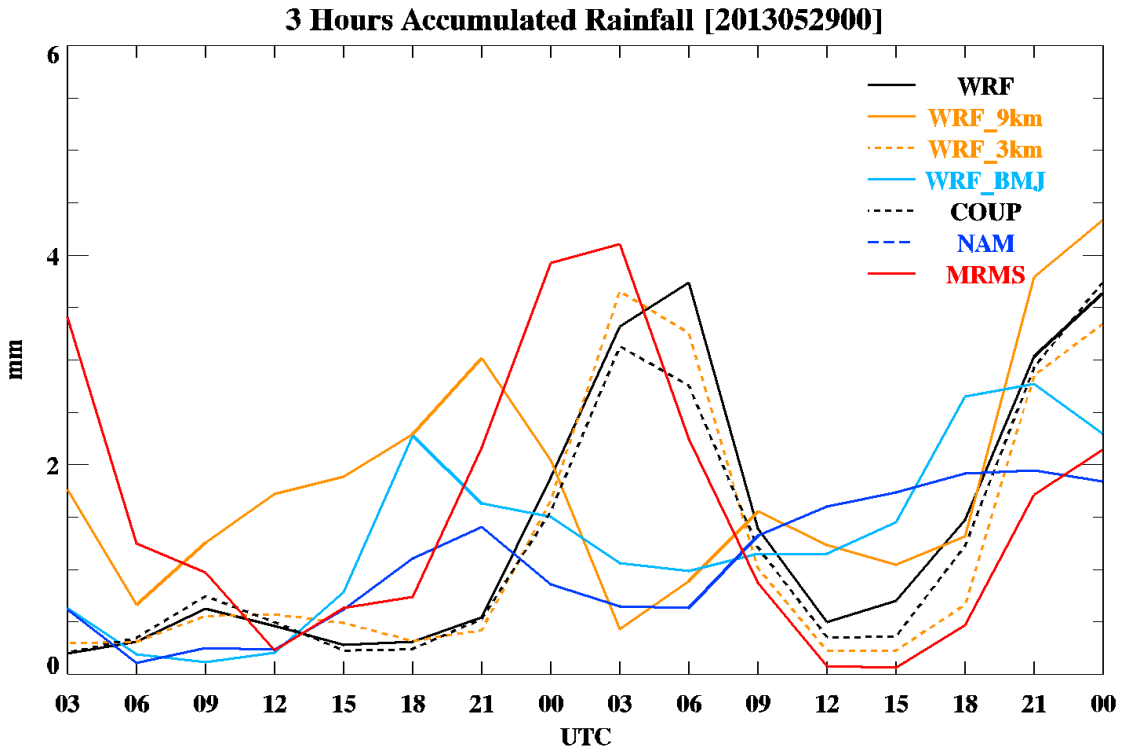


Figure 10: Domain averaged three hours rainfall accumulation for NU-WRF sensitivity runs. The result is from forecast cycle initiated at 00 UTC on May 29th, 2013.

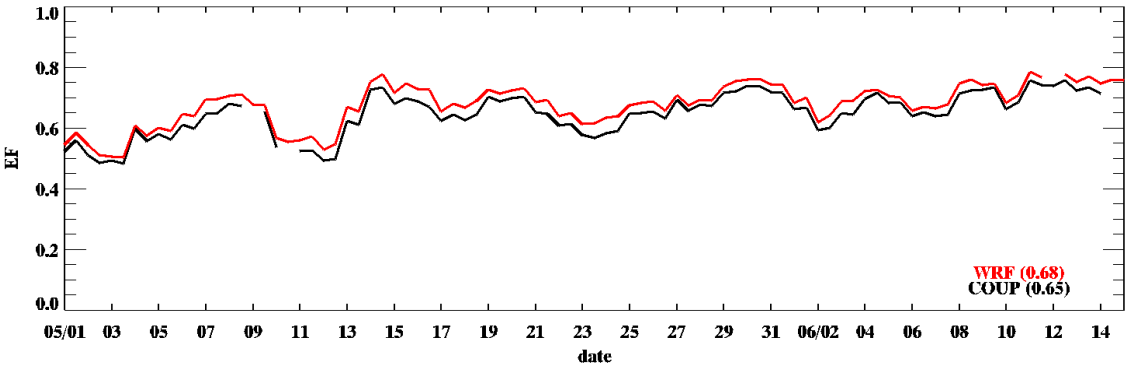


Figure 11: Daily averaged (7 am to 6 pm) evaporative fraction from May 1st to June 15th 2013.

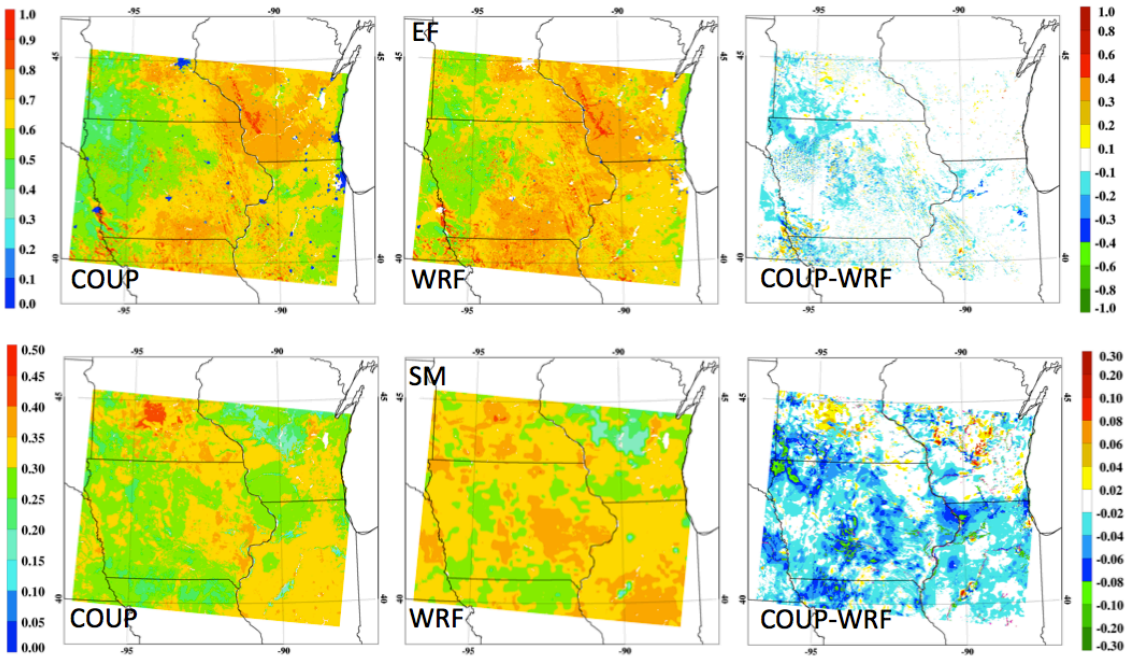


Figure 12: Evaporative fraction at local noon (17 UTC on May 19th) and top soil moisture (0-10 cm) at model initialization, which is at 12 UTC on May 19th, 2013.

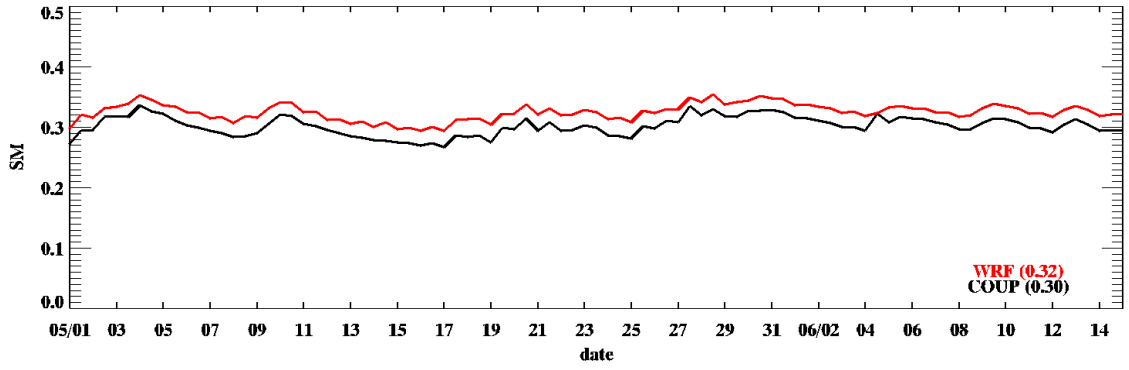


Figure 13: Same as Figure 11, except for daily averaged top soil moisture (0-10 cm).